PB 174707

ENVIRONMENTAL HEALTH SERIES
Air Pollution

SELECTION AND TRAINING OF JUDGES FOR SENSORY EVALUATION OF THE INTENSITY AND CHARACTER OF DIESEL EXHAUST ODORS

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service

ARCHIVE COPY

SELECTION AND TRAINING OF JUDGES FOR SENSORY EVALUATION OF THE INTENSITY AND CHARACTER OF DIESEL EXHAUST ODORS

Amos Turk, Ph. D.

Professor, Department of Chemistry
The City College of the City University of New York

Prepared under Public Health Service
Contract No. PH 27 66 96

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Bureau of Disease Prevention and Environmental Control
National Center for Air Pollution Control
Cincinnati, Ohio

1967

The ENVIRONMENTAL HEALTH SERIES of reports was established to report the results of scientific and engineering studies of man's environment: The community, whether urban, suburban, or rural, where he lives, works, and plays; the air, water, and earth he uses and re-uses; and the wastes he produces and must dispose of in a way that preserves these natural resources. This SERIES of reports provides for professional users a central source of information on the intramural research activities of programs and Centers within the Public Health Service, and on their cooperative activities with state and local agencies, research institutions, and industrial organizations. The general subject area of each report is indicated by the two letters that appear in the publication number; the indicators are

AP — Air Pollution
AH — Arctic Health
EE — Environmental Engineering
FP — Food Protection
OH — Occupational Health
RH — Radiological Health
WP — Water Supply and Pollution Control

Triplicate tear-out abstract cards are provided with reports in the SERIES to facilitate information retrieval. Space is provided on the cards for the user's accession number and additional key words.

Public Health Service Publication No. 999-AP-32

FOREWORD

Exhaust gases emitted by diesel engines are characterized by odors that are offensive in varying degrees to many members of the general public. The increasing use of diesel-powered trucks and buses in urban environments has resulted in widespread public awareness of the diesel exhaust odor problem.

It has not been established at this time whether any health hazards are involved; nevertheless, because the odors are unpleasant and irritating to people, studies of their origin and possible elimination are a necessary part of the over-all program of automotive air pollution research conducted by the National Center for Air Pollution Control, U. S. Public Health Service.

Several research organizations, working under PHS sponsorship, are conducting programs of engine testing and chemical analysis designed to establish the identity and relative concentrations of all products of incomplete combustion present in diesel engine exhaust.

To correlate these chemical analyses with the presence of odors, one must devise means of rating the odors, both as to quality and intensity, on some sort of numerical scale. Satisfactory correlations, once accomplished, would make it possible to evolve readily enforceable Federal standards for control of diesel exhaust odors — standards based on chemical composition rather than odor ratings and hence less subject to human error or bias.

This report, prepared under contract for the Public Health Service, outlines the development of training methods and chemical odor standards by which human panelists can measure the quality and intensity of diesel exhaust odors. These techniques are being applied in all of the PHS-sponsored research involving air pollution by diesel engines.

CONTENTS

Page

Abstract vii

Introduction 1

Environmental Conditions for Sensory Testing 3

Selection of Judges for Diesel Exhaust Odor Studies 5

The Training of Judges for Diesel Exhaust Odor Studies 15

Acknowledgments 19

Appendix A. Concentrations of Odorous Vapors in Test Chambers 21

Appendix B. Sensory Standards for Diesel Exhaust Odor Study 27

Appendix C. Statistical Derivations 35

ABSTRACT

Exhaust gases emitted by diesel engines are characterized by offensive odors. These odors must be rated numerically by human judges with the ultimate objectives of (1) correlating such ratings with the chemical composition of diesel exhaust and (2) establishing Federal standards for control of diesel exhaust odors.

Judges are selected on the basis of (1) their ability to distinguish among odors of different intensities and qualities, and (2) their behavior in subjective testing environments. The chosen judges are then trained to improve their performance in odor discrimination, to become familiar with diesel exhaust odor, and to rate the odor in terms of intensity and quality standards that are provided to them for reference. These standards comprise a scale of overall odor intensity, and four odor quality scales that correspond to the descriptions "burnt/smoky," "oily," "pungent/acid," and "aldehydic/aromatic."

Appendices describe (A) the theoretical basis for air purification requirements in test chambers for odor studies, (B) the composition and makeup of the diesel odor standards, and (C) the mathematical derivations of the statistical procedures.

INTRODUCTION

When a person is exposed to diesel exhaust on the street or highway, he may sense odor, chemical irritation in the nose or eyes, sound, velocity pressure of moving air, temperature gradients, impaction of small particles, and the sight of smoke, moving vehicles, and other associated actions. When the aggregate of such experience becomes sufficiently unpleasant, the threshold for individual action is exceeded and the affected person may voice a complaint or react in some other relevant manner.

Odor is undoubtedly the prime sensory attribute of diesel exhaust under the typical circumstances of human exposure. The sensory evaluation of diesel exhaust odor under different conditions of dilution, engine type, engine operation, fuel, fuel additives, and exhaust control devices will therefore provide part of the basis for specifying permissible conditions of exposure.

The sensory evaluations will be made by human judges. Taken as a group to increase the precision of the evaluations, the judges constitute a sensory odor panel. In serving on the panel, the individuals will be called upon to rank diesel exhaust samples to which they are exposed according to odor intensity and according to the quality, character, or type of odor. Depending on dilution of the exhaust, some samples will be chemically irritating (in the sense of "pungent") and odorous at the same time, even though the senses of irritation and odor are physiologically distinct. In some cases, therefore, the evaluation continuum will, in effect, comprise both types of sensations.

The judges in these tasks express their evaluations of intensity and quality in a quantitative way; they do not assume the role of individuals who express their personal preferences of "like" or "dislike." The following general requirements are imposed for the selection and performance of such judges.

1. They must have satisfactory sensory ability to distinguish among odors of different intensities and to discriminate among different odor qualities.

2. They must be emotionally receptive to tasks that involve quantitative and discriminatory judgments without expressions of preference.

3. They must be trained for the specific tasks to be accomplished.

4. A non-distracting environment must be provided for the sensory tests.

5. Standards should be established to provide a quantitative basis for sensory measurements and to provide replicable anchoring points that will facilitate interlaboratory comparisons.

ENVIRONMENTAL CONDITIONS FOR SENSORY TESTING

A space for sensory testing should be free of competing distractions. In-and-out visits and socialization by non-participants should be restricted, except in demonstration exercises. The area for preparation and coding of test samples or make-up and dilution of exhaust gas streams should be separated from the panel members and not visible to them. Extraneous sounds, especially the sounds of operation of test vehicles, should be inaudible during the tests. Color schemes should be neutral and not striking.

The test area should be provided with some means for space odor control, especially for control of the odors introduced from the samples being tested. This may be accomplished either by local exhaust of vapors from the test samples, or by general purification of the room air through a recirculating device, preferably an activated carbon unit, or by some combination of both methods. If a considerable volume of air containing odorous diluted vehicle exhaust is spilled into the space during the tests, then the effectiveness of odor control by general recirculation through an air purifying device will be limited. The nature of this limitation is described in detail in Appendix A.

When the tests are to be conducted and reported on an individual basis without communication between judges, it is desirable to have separate booths or enclosures for the individual panel members. When the test is to be cooperative, with discussion and comparisons among the judges, then a conference table setup is convenient.

The panel moderator or chairman should be able to communicate readily and conveniently with all of the judges. The general atmosphere should be comfortable and relaxing, but also should encourage the judges to be attentive and serious.

Several simple rules for conduct should be imposed. No smoking, eating, or drinking that is not associated with tests should be allowed in the test area. Women should be discouraged from the use of perfumes or perfumed cosmetics just prior to a test session. A period of about one-half hour or more should be allowed after smoking, eating a meal, or drinking coffee, before participation in a test exercise.

BLANK PAGE

SELECTION OF JUDGES FOR DIESEL EXHAUST ODOR STUDIES

There is no "magic number" of panel members. Many panels number five to fifteen judges, and a number close to ten is probably suitable for diesel exhaust sensory work. A large number of panelists should be available for testing so that no interruptions need be caused by temporary absenteeism or personnel turnover. The panel members should not be divided, however, into "regulars" and "standbys." Some rotation scheme should be set up so that no trained panel member is allowed to go "stale."

The major tasks that the odor judges will be expected to perform are:

1. To judge the relative intensity of diesel exhaust odors at different dilutions.
2. To discriminate among the different qualities of diesel exhaust odors.
3. To combine (1) and (2) to give a composite profile.

The judges will be expected to follow instructions, and at the same time to render independent judgments reflecting their own sensations and report their findings.

Four main tests for selection of judges are recommended:

1. The triangle test.
2. The intensity rating test.
3. The multi-component odor identification test.
4. Demonstration of satisfactory test behavior.

These tests are described in detail in the following sections. Statistical derivations of these tests are presented in Appendix C.

THE TRIANGLE TEST

Three test samples are presented at the same time. Two are identical, the third is different. The candidate is requested to identify the different or "odd" sample.

In dealing with untrained candidates, the moderator should avoid strange or objectionable materials. It is convenient to use dilute aqueous solutions of food flavors. These are easy to administer, and they product a minimum of initiation "shock" to the candidate. The solutions may be smelled or tasted; in both cases, olfactory discrimination is actually being used to make the identification. Materials that are recommended are:

Flavorant	Approximate Concentrations
Vanilla extract Lemon extract Pineapple extract Almond extract Rum extract Rose extract Mint extract	To one quart of water add 1 or 2 drops of required extract to achieve a detection threshold at a level such that about 75% of the triangle scores will be correct answers.

The samples should be presented at ambient temperature in small throw-away paper cups.

Instructions are given as follows: "The object of this test is to distinguish differences between food flavors that are very weak in intensity. The samples given to you consist of one or two drops of a common food extract, like vanilla, almond, or mint, in a quantity of pure drinking water. You may smell or taste the samples, or both. Each test consists of three samples, two of which are the same and one of which is different. Your task is to pick out the different or odd one. You do not have to identify the flavor, just choose which one is different from the other two. Enter on your score sheet the code number of the sample which you choose as the odd one."

Present a trial test to familiarize the candidate with the procedure. This trial should be conducted at a higher level of intensity so that the distinction is easy. Do not ask for a score. State the correct answer verbally to the candidate. Answer any questions regarding procedure. The purpose of this trial is to concentrate most of the learning into the first experiment and to make the subsequent tests independent of each other to a greater degree. Now administer and score five triangle tests.

Select the candidates on the following basis:

List the candidates and their scores (number of correct answers) in order, highest score first.

Refer to Table I to determine which differences between scores are significant.

Decide how many panel members you should select from the candidates available. Determine from the table at what levels there are statistically significant differences between candidates and make your selection accordingly. For these rough screening operations, it is recommended that confidence limits be set at about 25 or 30%, rather than the usual 5%. This will have the effect of initial rejection of most of the candidates that are likely to be unsuitable.

Example I. There are ten candidates, who take five triangle tests each. We need about five or six panel members from among these candidates. Their scores are:

Candidate	Number of correct answers
A	5
B	2
C	3
D	0
E	0
F	2
G	3
H	5
I	5
J	4

TABLE I — SIGNIFICANCE OF DIFFERENCES AMONG SCORES IN TRIANGLE TESTS

Number of tests	Difference between two candidates in number of correct answers	Probability that at least this difference could have been obtained by chance
1	0	100%
	1	44
2	0	100%
	1	59
	2	10
3	0	100%
	1	66
	2	19
	3	2
4	0	100%
	1	71
	2	25
	3	5
	4	0.5
5	0	100%
	1	74
	2	31
	3	9
	4	2
	5	0.1

Step 1. Choose an acceptable significance level. In this case we will choose 0.20-0.30 (70-80% confidence levels).

Step 2. List the candidates in order of the number of correct answers they gave.

Candidates	Number of correct answers
A, H, I	5
J	4
C, G	3
B, F	2
D, E	0

Step 3. Check Table 1 to determine how much of a difference in the scores of two candidates is needed to give confidence that they are significantly different from each other. In this example, the number of tests equals 5, therefore, a difference of 2 or more in score indicates a significant difference in performance.

Step 4. Choose the candidates required. Starting from the top, A, H, and I have a score of 5. But the scores of candidates A, H, I, and J differ by less than 2; therefore this difference is not significant, and we can consider these candidates to be "equal." If we need five or six panelists and have no other candidates, we must add C and G to the panel. (J, C, and G can be considered "equal.") The panel will then consist of the first six: A, H, I, J, C, and G.

Note that we could establish higher confidences in our selections by giving more tests, say 20 or 25 instead of 5. This is not recommended because the extra effort can be more profitably expended in the screening of additional candidates.

THE INTENSITY RATING TEST

A series of dilutions of an odorant in an odorless diluent is set up. One sample is removed from the series. The candidate is asked to replace it according to its odor intensity in the position from which it was taken.

For use in selecting candidates, the odorant should be fairly strong when it is in pure form, and not toxic, unpleasant, or strange. Any of a number of fruity or fragrant odors are acceptable. Some possibilities are amyl acetate, eucalyptol, oil of wintergreen, or heptaldehyde.

In the example that follows the odorant is amyl acetate in propylene glycol (low-odor perfumer's grade). Amber 2-ounce glass bottles with plastic screw caps are used.

Bottle No.	Procedure	Concentration (fraction by volume of amyl acetate in the solution)
1	Add 10 ml amyl acetate.	1
2	Add 10 ml amyl acetate plus 10 ml propylene glycol. Mix. Using a 10 ml pipet, remove 10 ml of the mixture and transfer to bottle 3.	½
3	Contains 10 ml of material from bottle 2. Add 10 ml propylene glycol. Mix. Remove 10 ml of mixture and transfer to bottle 4.	¼
4	Contains 10 ml of material from bottle 3. Add 10 ml propylene glycol. Mix. Remove 10 ml of mixture and transfer to bottle 5.	⅛

Continue this procedure until the mixture has no detectable odor of amyl acetate. The total number of bottles thus prepared will be approximately 20.

If the bottle number is designated **n**, then the concentration in any bottle equals 2^{1-n}.

The bottles must appear identical and must be consecutively numbered so that the numbers are not visible to the candidates. Labels can be put on the bottoms of the bottles or on the face of the bottle if the label is obscured by wrapping with aluminum foil during the test. The bottles are lined up in sequence in front of the candidate.

The following instructions are given: "The 20 bottles lined up in front of you all contain solutions of amyl acetate, which is a synthetic banana oil. They differ from each other in odor strength, the most intense odor is on the left, and the intensity gradually gets less from bottle to bottle toward the right. The last bottle on your right has so little banana oil odor that it may not be detectable at all. In the test you are about to perform one of these bottles will be removed from the series, from a position unknown to you. The task will be to replace it in the proper location in the series, the location from which it was taken. If it is replaced in its proper position, it will smell stronger than the bottle on its right and weaker than the bottle on its left. Proceed as follows. First familiarize yourself with the odors of the bottles in the series. Start from the right (number 20) remove the cap, sniff gently and recap. Then do the same for about every other bottle going to your left. Remember, the odors will be getting stronger toward the left and you will fatigue your sense of smell temporarily

if you sniff too long at the more intense odors. You need not smell every bottle in the series at this stage of the test." (After a brief time for familiarization . . .) "Now leave the room." (The tester removes one bottle, and rearranges the others to obscure the gap. The removed bottle is placed in front of the remaining 19). "Now come back into the room and replace the bottle in the location where it belongs in the series."

Statistical scoring procedure: The **same** specified test program, in randomized order, must be used for each candidate. This program consists of four tests, in which the following bottle numbers are successively removed in random order: 3, 8, 12, and 16. This presupposes a dilution series of at least 20 bottles. The highest test bottle number must be 4 less than the number of bottles in the series. If the total number of bottles is, say 17, the test sequence could be 3, 7, 10, 13.

Step 1. Score as follows:

Positions removed from correct location	Score (higher number is worse score)
0	0
±1	1
±2	4
±3	9
±4 or more	16

Step 2. Refer to Table 2 to determine which differences between scores are significant.

TABLE 2 — SIGNIFICANCE OF DIFFERENCES AMONG SCORES IN INTENSITY RATING TEST

Differences between scores of any two candidates	Probability that at least this difference could have been obtained by chance
18	30%
20	25%
23	20%
25	15%
29	10%
34	5%
44	1%
56	1/10%

Example 1. A candidate performs as follows:

Missing bottle	Replaced in position
8	7
12	12
3	5
16	6

What is the candidate's score?

Bottle Number	Positions removed from correct location	Score
8	8—7 = 1	1
12	12—12 = 0	0
3	3—5 = —2	4
16	16—6 = 10	16
		Total score 21

Example 2. Two candidates show the following performances. Is one better than the other? What confidence is there in the selection?

Candidate A

Bottle Missing	Replaced in position
3	1
16	12
8	12
12	15

Candidate B

16	13
12	15
3	3
8	10

A's score

	Score
3—1 = 2	4
16—12 = 4	16
8—13 = —5	16
12—15 = —3	9
	45

B's score

	Score
16—13 = 3	9
12—15 = —3	9
3—3 = 0	0
8—10 = —2	4
	22

A—B = 45—22 = 23

Therefore **B** scored better than **A**. From the significance table (Table 2), this could have been due to chance alone 20% of the time. Therefore, we have a confidence of 80% in our selection of **B** over **A**. If we demand a confidence of, say, 90% before making a decision, then we cannot rule out **A** as a possible panel member before further testing.

THE MULTI-COMPONENT ODOR IDENTIFICATION TEST

This test presents three mixtures to the candidate. These mixtures contain, in sequence, 2, 3, and 4 odors out of a possible total of 8 known standards. The candidate is told how many components to look for, and is asked to identify them.

The following materials are recommended as standards:

1. Oil of cade (burnt)
2. Cassia (cinnamon)
3. Eucalyptus
4. Amyl acetate (banana)
5. Clove oil
6. Orange oil
7. Almond oil (benzaldehyde)
8. Vanillin or vanilla extract

Instructions are given as follows: "You have in front of you eight labeled common odors. They are (state which). These will be your reference standards and you may smell them whenever you wish. Now, you will be given an unknown odor sample. This sample contains a mixture of two of the standard odors. Write down on your score sheet which two odors are present in the test sample." When this test is finished, repeat the instructions but state that the new sample contains a mixture of three standards. Then repeat the instructions again for the third test explaining that the sample contains a mixture of four odors.

The multi-component odor identification test is scored as follows:

Step 1. Administer the **same** test sequence, consisting of three tests, to each candidate.

Test 1. Two components

Test 2. Three components

Test 3. Four components

Step 2. Calculate the score of each individual candidate according to the following procedure.

$$\text{Total score} = \left(\begin{array}{c}\text{Number correct}\\ \text{in test 1}\end{array}\right)^2 + \left(\begin{array}{c}\text{Number correct}\\ \text{in test 2}\end{array}\right)^2 + \left(\begin{array}{c}\text{Number correct}\\ \text{in test 3}\end{array}\right)^2$$

Step 3. Refer to Table 3 to determine which differences between scores are significant.

TABLE 3 — SIGNIFICANCE OF DIFFERENCES BETWEEN SCORES IN MULTI-COMPONENT ODOR IDENTIFICATION TEST

Difference between scores of any two candidates	Probability that at least this difference could have been obtained by chance
6	25%
7	20%
8	15%
9	10%
11	5%
15	1%

Example 1. Candidates A and B scored as follows in the multi-component odor identification test. Are the scores significantly different?

Test	Number Correct	
	Candidate A	Candidate B
1	2	0
2	3	1
3	1	4

Score of A $= 2^2 + 3^2 + 1^2 = 14$

Score of B $= 0^2 + 1^2 + 4^2 = 17$

Difference = 3. The probability that this difference could have been obtained by chance is greater than 25%. Therefore, we do not attach any significance to this difference, and the two candidates should not be differentiated from each other.

Example 2. Choose the best four candidates from the following:

Test	Number correct obtained by each candidate.								
	A	B	C	D	E	F	G	H	I
1	0	2	1	1	1	0	2	0	2
2	3	3	2	0	1	0	1	1	3
3	4	2	4	1	3	4	0	2	3

The scores are computed on the basis of the sum of the squares of the correct answer and are then listed in order of magnitude. The order is

Candidate	Score
A	25
I	22
C	21
B	17
F	16
E	11
G, H	5

Now, referring to Table 3, we can say with a confidence of 80% that any candidates whose scores differ by no more than 7 points are considered "equal." By this criterion of confidence, candidates A, I, and C could be considered to be equal and could be selected for the panel. Candidate B differs from Candidate A by more than 7 points and is therefore ruled out at this confidence level. At a confidence level of 90%, however, which implies a lower probability (10%) that the observed differences of scores could have been obtained by chance, we could have selected all of the candidates except the last three.

DEMONSTRATION OF SATISFACTORY TEST BEHAVIOR

During the phases of candidate selection, the panel leader or moderator should be attentive to various aspects of the candidates' test behavior. These aspects are:

Speed. The best behavior is purposeful and deliberate, neither excessively hasty nor slow.

Interest level. The candidate should feel challenged and motivated. Candidates who find the work distasteful or uncomfortable should not be selected.

Domination. In group testing, the candidate should be helpful when asked, but should not try to push his or her opinions on others.

Independence. The candidate should be willing to consider the suggestions of others, but should not be influenced to change score against his own judgment.

Honesty. Candidates who try (successfully or not) to decode labels or peek under bottles should be rejected.

The panel leader should observe the candidates' behavior in these matters carefully but unobtrusively. No specific numerical scoring system is recommended; the leader should rely on his own judgment with regard to the sensitive question of whether to reject, on the basis of poor test behavior, any otherwise acceptable candidates.

THE TRAINING OF JUDGES FOR DIESEL EXHAUST ODOR STUDIES

Training should be scheduled according to the following procedure.

1. Improve the performance of the judges in the same tests that were administered for panel selection. (Estimated time: 3-4 days).

2. Expose the judges to diesel exhaust under the conditions to be used in the testing, under conditions of random ambient air dilution, to diesel fuel, and to components of diesel exhaust condensate if such are available. Discuss the four diesel odor descriptors (see Appendix 3) during each of these exposures. (Estimated time: 2-3 days).

3. Introduce the judges to the kit of diesel odor reference standards. Explain how the standards will be used in measuring intensity and quality of diesel exhaust odor. Train the panel members in recognizing individual samples in the kit and in recognizing multicomponent samples. Establish minimum requirements for these tasks. During this training, intersperse exercises that relate the diesel quality standards to the sensory quality of diesel exhausts. (Estimated time: 1-2 weeks).

4. Eliminate panel members as necessary in phases A, B, and C for reasons of poor performance or motivation, emotional problems, excessive domination or dependence, or other difficulties apparent to the moderator.

5. The panel may now operate according to whatever experimental design is adopted as the original research program. The continued conduct of the work will, in effect, continue the training of the panel. The panel members should be checked periodically to insure that the level of the performance with the kit components obtained in phase 3 is maintained.

Some of these phases of panel training are considered in more detail in the following sections.

INITIAL IMPROVEMENT OF PERFORMANCE

The tests used for the selection of panel members should be repeated, with one important difference. In training, no errors are ignored. As soon as a trainee or group of trainees scores a test, the correct answers are disclosed; the trainee repeats the exercise with attention focused on elimination of the error, with the help of the leader.

Example 1.

Trainee: "I detect eucalyptus, banana, clove, and cinnamon in this mixture."

Leader: "The mixture contains eucalyptus, banana, almond, and cinnamon. That means you reversed clove and almond. Refer to the almond and clove standards and recheck them and then smell your mixture again."

(Note — The leader must not be deprecatory. Say "reversed" to express the facts, not "confused" or "made an error" or "goofed" to express disapproval.)

Trainee: "Well, that almond is overpowered by the other odors. It didn't come through for a while."

Leader: "That's all right, it may happen that way. A four-component mixture is a difficult task. Focus your attention on one component at a time, take only short sniffs so you don't get fatigued, and refer to the standards whenever you have to."

During training, allow trainees to work together and help each other by suggestions and exchange of comments and samples. Keep records of the trainees' progress.

Training should be continued until there is a noticeable levelling-off in performance. Prolonged continuation of training on known mixtures is wasteful, because the diesel work will involve new materials and, of necessity, new training.

An important aspect of training, and, later, of panel utilization is the question of whether the panel should work on an individual or on a group basis. The individual basis eliminates social influences; each panelist's score is purely his own judgment. The group basis allows for cooperative suggestions that will call a panelist's attention to something that he himself senses only after his attention is properly focused. The decision between the two methods is best made on a statistical basis. It will be appropriate to use the t-test to determine whether or not the difference between the average panel scores obtained on an individual and on a group basis is significant.

EXPOSURE TO DIESEL EXHAUST

The panel members must be exposed to diesel exhaust under various conditions that are likely to be experienced in the proposed testing program and also under typical conditions expected on the street or highway. The panelists should also be exposed to diesel fuel and/or fuel components, diesel exhaust condensate, and other relevant odor sources. Explain to the panel members that the quality of an odor can be described in terms of quality components. The explanation can be an expansion of remarks such as the following:

"You have identified the components of mixtures of two, three, or four odorants in the tests you have been conducting during the last few days. This identification is also a description of odor quality. You may say that the odor of a given sample is a mixture of lemon and vanilla. This statement describes the quality of the odor mixture,

just as the same statement would be a description of flavor if the sample were a piece of cake. We are going to describe diesel exhaust in terms of four qualities:

1. Burnt/smoky. This is the quality in which diesel exhaust is related to the odors of other products of burning or combustion.

2. Oily. This quality is the oiliness related to the presence of the odor of the heavy components of unburned fuel.

3. Pungent/acid. This is the quality associated with pungency in high concentration, gradually changing to an acid or sour quality with greater dilution.

4. Aldehydic/aromatic. This is the quality related to what may be thought of as the "fragrant" aspect of diesel exhaust. It may also be thought of, in most instances, as that quality *not* represented by smokiness, oiliness, or pungency."

LEARNING THE DIESEL ODOR STANDARDS

(Refer to Appendix B for details of the composition and preparation of the diesel odor kit.)

The following introductory remarks may be appropriate:

"You will be given a kit of standard odors to which you will refer in making your odor judgments. Standard odors D-1 through D-12 represent the intensity levels of diesel exhaust odor, without regard to quality. Standards B-1 to B-4 represent four intensity levels of burnt odor; O-1 to O-4 represent intensity levels of oily odor. (The moderator at this point should explain the entire kit.) Before using these standard odors to measure diesel exhaust, it will be necessary for each of you to learn the various standards and to get some familiarity with mixtures of the standards."

The training for recognition of odor standards should now proceed along the following lines.

1. The judges should be trained in the intensity levels of the D-1 to D-12 series. Judges should be able to identify any unknown sample to within $\pm$ 1 or 2 intensity levels.

2. The judges should learn each of the quality components at the "extreme" concentration (B-4, O-4, P-4, and A-4). They should then learn to identify each intensity of each quality, 16 bottles in all. They should be able to identify all unknowns as to quality and intensity about 90% of the time.

3. The judges should learn to identify two- and three-component mixtures, until performance levels off. They should reach a level at which they identify mixtures correctly about 80% of the time. The compositions of standard mixtures are shown in Table 4 and 5. Probabilities of obtaining correct identifications by chance are shown in Table 6.

TABLE 4 — COMPOSITION OF COMPONENTS FOR ODOR STANDARD MIXTURES

Quality	Code (t = training)	Composition
Burnt	B-t	B-4 (See Appendix 3)
Oily	O-t	100% octylbenzene
Pungent	P-t	P-4
Aldehydic	A-t	1% A in mineral oil

TABLE 5 — COMPOSITION OF ODOR STANDARD MIXTURES FOR TRAINING

Number of Components	Code (t = training)	Composition
2	AB-t AO-t AP-t BO-t BP-t OP-t	1 : 1 mixtures of the solutions of Table 4. About ½ to 1 ml of solution is placed in a wad of cotton in a plastic squeeze bottle.
3	ABO-t ABP-t AOP-t BOP-t	1 : 1 : 1 mixtures as above.
4	ABOP-t	1 : 1 : 1 : 1 mixture as above.

TABLE 6 — PROBABILITY OF CORRECT IDENTIFICATION OF TRAINING STANDARDS BY CHANCE ALONE

Number of components stated to be in mixture	Probability per test
1	1/4
2	1/6
3	1/4
Not disclosed	1/15

ACKNOWLEDGMENTS

The following chemists participated in the development of the odor standards: Louis Reckner, Robert Squires, and Z. Tomaras, all of the Scott Research Laboratories; Stanley Mehlman and Elizabeth Wolf, both of the City University of New York; and Jonathan Turk, of Brown University.

The statistical calculations were derived by Janet T. Wittes of the Department of Statistics, Harvard University.

APPENDIX A: CONCENTRATIONS OF ODOROUS VAPORS IN TEST CHAMBERS

Odor test chambers are used either for evaluation of odor-reducing devices or to provide odor-free environments in which a jury can measure the odors of materials, products, or foods. It is important to consider the factors that determine the changing or equilibrium concentrations in such chambers.

Processes that tend to increase the concentration of odorous vapors are:

(a) the generation of vapor within (or injection into) the space,

(b) the introduction of vapor by replacement of chamber air by ventilation or infiltration with outdoor air of higher vapor concentration.

Processes that tend to decrease the concentration of odorous vapor are:

(a) the treatment of the chamber air by a vapor-reducing device (e.g., activated carbon recirculator),

(b) the removal of vapor by replacement of chamber air by ventilation or infiltration with outdoor air of lower vapor concentration.

The concentration of odorous vapor in a chamber will approach an equilibrium value in which the rates of vapor-reducing and vapor-increasing processes are equal. If it is assumed that air introduced into the chamber by ventilation, infiltration, or recirculation through a treatment device is completely and instantaneously mixed with the chamber air, then the concentration of vapors at any time and at equilibrium are given by the general equations:

$$C = C_o e^{-(Q_i + EQ_r)t/V} + \left(\frac{C_i Q_i + G}{Q_i + EQ_r}\right)\left[1 - e^{-(Q_i + EQ_r)t/V}\right]$$

$$C_\infty = \frac{C_i Q_i + G}{Q_i + EQ_r}$$

SPECIAL CASES OF THE GENERAL EQUATIONS

Consider the following possibilities:

(a) Ventilation air is pure. ($C_i = 0$)

(b) No vapor is being generated or injected. ($G = 0$)

(c) The vapor reducing device is 100% efficient. ($E = 1$ and $C_r = 0$)

(d) The chamber is originally pure. ($C_o = 0$)

(e) The room is tight. ($Q_i = 0$)

(f) Combination of (a) and (b)

(g) Combination of (a) and (c)

(h) Combination of (b) and (e)

(i) Combination of (a), (b), and (c)

For each possibility (a) through (i), the equation for the vapor concentration at any time (C) or at equilibrium (C_∞) is derived directly from the general equations:

(a) $C_i = 0$

$$C = C_o e^{-(Q_i + EQ_r)t/V} + \frac{G}{(Q_i + EQ_r)} \left[1 - e^{-(Q_i + EQ_r)t/V}\right]$$

$$C_\infty = \frac{G}{(Q_i + EQ_r)}$$

(b) $G = 0$

$$C = C_o e^{-(Q_i + EQ_r)t/V} + \frac{C_i Q_i}{(Q_i + EQ_r)} \left[1 - e^{-(Q_i + EQ_r)t/V}\right]$$

and

$$C_\infty = \frac{C_i Q_i}{(Q_i + EQ_r)}$$

(c) $E = 1$ and $C_r = 0$

$$C = C_o e^{-(Q_i + Q_r)t/V} + \left(\frac{C_i Q_i + G}{Q_i + Q_r}\right) \left[1 - e^{-(Q_i + Q_r)t/V}\right]$$

and

$$C_\infty = \frac{C_i Q_i + G}{(Q_i + Q_r)}$$

(d) $C_o = 0$

$$C = \frac{C_i Q_i + G}{Q_i + EQ_r} \left[1 - e^{-(Q_i + EQ_r)t/V}\right]$$

Note that the concentration rises to the equilibrium value expressed by the general equation.

(e) $Q_i = 0$

$$C = C_o e^{-EQ_r t/V} + \frac{G}{EQ_r}\left(1 - e^{-EQ_r t/V}\right)$$

and

$$C_\infty = G/EQ_r$$

(f) Combination of (a) and (b). C_i and $G = 0$

$$C = C_o e^{-(Q_i + EQ_r)t/V}$$

and

$$C_\infty = 0$$

(g) Combination of (a) and (c). $C_i = 0$, $E = 1$, and $C_r = 0$

$$C = C_o e^{-(Q_i + Q_r)t/V} + \frac{G}{(Q_i + Q_r)}\left(1 - e^{-(Q_i + Q_r)t/V}\right)$$

and

$$C_\infty = \frac{G}{Q_i + Q_r}$$

(h) Combination of (b) and (e). $G = 0$, and $Q_i = 0$

$$C = C_o e^{-EQ_r t/V}$$

and

$$C_\infty = 0$$

(i) Combination of (a), (b), and (c). $C_i = 0$, $G = 0$, $E = 1$, and $C_r = 0$

$$C = C_o e^{-(Q_i + Q_r)t/V}$$

and

$$C_\infty = 0$$

EXPRESSION IN TERMS OF AIR CHANGES

An air change is the addition to the chamber of a volume of air equal to the volume of the chamber. Then the number (N) of air changes (dimensionless) per unit time is given by:

$$N/t = Q/V, \text{ or}$$
$$N = Qt/V$$

MIXING FACTOR

In any expression of the general form $e^{-Qt/V}$ or e^{-N}, a mixing factor, m, may be applied to account for the fact that dilution of air is not instantaneous, and that concentration fall-off rates are actually smaller than the ideal values given by the equations developed here. Brief[2] suggests that m commonly ranges between 1/3 and 1/10.

SENSORY TESTING

The efficiency of a vapor-reducing device may be measured by sensory methods that obviate the need for determination of material concentrations. If we select a sealed chamber in which no odor is being generated, then the ratio of concentrations C_1/C_2 corresponding to any two times t_1 and t_2 during the operation of the device is

$$C_1/C_2 = e^{(EQ_rm/V)\ (t_2 - t_1)}$$

It is possible to measure the ratio of two supra-threshold concentrations of odorous vapor C_1/C_2 by either a dilution or a matching technique. Let

$$P = \frac{\text{Volume of air sample diluted to threshold}}{\text{Volume of original air sample}} = C/C_t$$

Then, the ratio of P values for any two concentrations C_1/C_2 is:

$$\frac{P_1}{P_2} = \frac{C_1/C_t}{C_2/C_t}, \text{ or}$$

$$P_1/P_2 = C_1/C_2$$

We may therefore write

$$P_1/P_2 = e^{(EQ_rm/V)\ (t_2 - t_1)}$$

and

$$\log_{10} P_1/P_2 = 0.434\ (EQ_rm/V)\ (t_2 - t_1)$$

Solving for E,

$$E = \frac{2.303 \log_{10} P_1/P_2}{(Q_rm/V)\ (t_2 - t_1)}$$

The latter equation tells us that for a room of known volume and air mixing characteristics, it is possible, by sensory determination odor in the room at different times, to measure the efficiency of a vapor-reduction device operating within the room.

Example: In a 1000-cu.-ft. sealed room, 50 cfm of air is recirculated by an air purifier of 60% efficiency. The mixing factor in the room is ⅓. Some diesel exhaust vapor is spilled into the room. How long must the air purifier operate if the vapor concentration in the room is to be reduced by 90%?

Solution: $C_2 = C_1 - 0.9C_1$, and $C_1/C_2 = 10$

$$t_2 - t_1 = \frac{2.303 \log C_1/C_2}{EQ_rm/V} = \frac{2.303 \log 10}{0.60 \times 50 \times ⅓/1000}$$

$$= 230 \text{ min., or } 3 \text{ hr. } 50 \text{ min.}$$

NOTATION

V = Volume of chamber
t = time
C = concentration of vapor in chamber at any time
C_0 = initial concentration of vapor in chamber
C_∞ = concentration of vapor in chamber at equilibrium
C_i = concentration of vapor in ventilation or infiltration air
C_r = concentration of vapor delivered by the air treatment device
E = efficiency of vapor reduction by the air treatment device
Q_i = volume rate of ventilation or infiltration
Q_r = volume rate of air delivery by the air treatment device
G = quantity rate of generation of vapor within (or injected into) chamber
N = number of air changes
m = mixing factor

REFERENCES

1. A. Turk, Measurements of Odorous Vapors In Test Chambers: Theoretical. ASHRAE J., Oct. 1963.

2. R. S. Brief, Simple Way to Determine Air Contaminants, Air Engineering, Vol. 2, p. 39 (1960).

APPENDIX B: SENSORY STANDARDS FOR DIESEL EXHAUST ODOR STUDY

This Appendix describes the composition and method of preparation of sensory standards for diesel exhaust odor study. The standards are designed to serve as sensory references for (a) the intensity of diesel exhaust odor, without regard to the quality of the odor, (b) the quality of diesel exhaust odor, expressed in terms of the intensities of each of four odor quality descriptors, and (c) the intensity of odor-modifying agents that are designed to improve diesel exhaust odor by odor masking or odor counteraction.

The intensities of the individual odor qualities, taken together, constitute a "quality-intensity profile," or "QI profile." The profile is **not** designed so that the sum of its intensity scores should be related to the overall diesel exhaust odor intensity.

Descriptors are used to designate diesel exhaust odor quality.

The descriptors are:

Name	Code
Burnt/smoky	B
Oily	O
Pungent/acid	P
Aldehydic/aromatic	A
Masking	M

BURNT/SMOKY (B)

Burnt quality or smokiness is a typical odor component of the products of combustion of organic matter. There is considerable variation in the quality of burnt odors, however, among different conditions of combustion, different materials being burned, and different states of molecular or aerosol aggregation of the airborne combustion products. Many of the primary combustion products are unstable and therefore unsuitable for use as sensory reference standards. The chemical makeup of materials that have a burnt odor includes products of decomposition and partial oxidation. Oil of cade (juniper tar) is included in the B standard because this oil has a typical burnt odor and is readily available from commercial sources. Guaiacol and carvacrol impart the phenolic odor component that is contributed to some extent by the oxidation of benzenoid aromatic matter. Acetylene dicarboxylic acid is a commercially available chemical whose odor somewhat resembles that of a dilute mixture of carbon suboxide in air. Carbon suboxide is a likely product of partial oxidation that contributes to burnt odor. Benzyl benzoate is an almost odorless diluent that solubilizes the other components.

Composition of B

Component	Percent by weight	
	including solvent	excluding solvent
Oil of cade (Juniper tar)	16.8	84.3
Guaiacol	0.2	1.0
Carvacrol	0.9	4.2
Acetylene dicarboxylic acid	2.1	10.5
Benzyl benzoate	80.0	(solvent/odorant, 4:1)
	100.0	

Intensity series (diluent: as noted)

		Concentration (B/B + diluent)			
Odor intensity	Code	Diluent	Fract.	Decimal	%
Slight	B-1	min. oil	1/720	0.00319	0.139
Moderate	B-2	min. oil	1/180	0.00556	0.556
Strong	B-3	min. oil	1/45	0.0222	2.22
Extreme	B-4	benz. ben.	1/4	0.25	25.

Procedure

Make up the stock mixture B.

To make B-4, mix 1 part of B with 3 parts of diluent.

To make B-3, mix 1 part of B-4 with 44 parts of diluent.

To make B-2, mix 1 part of B-3 with 3 parts of diluent.

To make B-1, mix 1 part of B-2 with 3 parts of diluent.

Notes: (a) The dilutions can be made on a volume basis without introducing any significant error.

(b) Stock solution B should be shaken before dilutions are made to bring into suspension any component of the oil of cade that might have settled.

(c) Note the change in diluent between B-4 and B-3.

OILY (O)

Oiliness is an odor quality generally associated with organic chemicals whose molecular structure is characterized by long saturated hydrocarbon chains. Among materials of plant origin such substances are typically esters of long-chain fatty acids. In diesel exhaust, oily quality is believed to be associated with the presence of high-boiling-point components of unburned fuel, in the boiling range approximately 300° C (1 atm). The most satisfactory standard found to represent this odor quality is n-octylbenzene.

Composition of O

n-Octylbenzene, 100%

Intensity series (diluent: mineral oil)

		Concentration, O/(O + diluent)			
Odor intensity	Code	Expon.	Fract.	Decimal	%
Slight	O-1	2^{-7}	1/128	0.0078	0.78
Moderate	O-2	2^{-5}	1/32	0.0313	3.13
Strong	O-3	2^{-3}	1/8	0.125	12.5
Extreme	O-4	2^{-1}	1/2	0.5	50.

Procedure

The stock material O is pure n-octylbenzene.

To make O-4, mix 1 part of O with 1 part of diluent.

To make O-3, mix 1 part of O-4 with 3 parts of diluent.

To make O-2, mix 1 part of O-3 with 3 parts of diluent.

To make O-1, mix 1 part of O-2 with 3 parts of diluent.

Note: (a) The dilutions can be made on a volume basis without introducing any significant error.

PUNGENT/ACID (P)

Diesel exhaust in high concentration can be perceived as an irritant. The word "irritant" is used here to denote a substance that can be detected by the common chemical sense, as distinguished from the specific olfactory sense. Such irritants are said to have a "pungent" quality. This type of common chemical sensation can coexist with odor. Thus, a concentrated mixture of butyric acid vapor in air is both pungent and odorous. As the mixture is diluted, the common chemical irritation diminishes and then disappears at concentrations at which odor still persists. The sensation is then no longer said to be pungent; instead, it is described as "acid," or "sour." The existence of pungent, acid components in diesel exhausts is evidenced by (a) the interpretation of infrared spectra by Scott Research Laboratories to indicate the presence of organic acid (carboxyl function), (b) the sensory identification of "sour" substances among the column chromatographic fractions of diesel exhaust obtained by Scott Research Laboratories, and (c) the pungency experienced by direct exposure to concentrated diesel exhaust.

The character of "sourness" or "acid odor" (as distinguished from pungency) varies among different sources. Butyric and valeric acid odors are characterized by sour odor qualities typical of rancidification of organic matter. An acidic odor quality more closely related to combustion products is associated with organic acids that have olefinic or acetylenic unsaturation.

Composition of P

Component	Percent by weight
Crotonic acid	3.3
Propiolic acid	2.9
Benzyl benzoate	93.8
	100.0

Intensity series (diluent: mineral oil)

		Concentration, P/(P + diluent)		
Odor intensity	Code	Fract.	Decimal	%
Slight	P-1	1/720	0.0014	0.14
Moderate	P-2	1/180	0.0055	0.55
Strong	P-3	1/60	0.0167	1.67
Extreme	P-4	1/20	0.05	5.0

Procedure

Make up the stock solution P. The mixture will have to be warmed slightly to bring the components into solution. Ailow the solution to cool to ambient temperature.

To make P-4, mix 1 part of P with 19 parts of diluent.
To make P-3, mix 1 part of P-4 with 2 parts of diluent.
To make P-2, mix 1 part of P-3 with 2 parts of diluent.
To make P-1, mix 1 part of P-2 with 3 parts of diluent.

Note: (a) The dilutions can be made on a volume basis without introducing any significant error.

(b) The components of P-4 may have to be warmed to bring them into solution.

ALDEHYDIC/AROMATIC (A)

Aldehydes are known to exist as components of diesel exhaust. Some of the column chromatographic fractions of diesel exhaust obtained by Scott Research Laboratories were characterized as having "sweet" or "spicy" odors. A mixture used at Scott Research Laboratories for setting up odor intensity ratings of diluted diesel exhaust contained heptaldehyde as the major component. These circumstances, taken together, support the selection of a quality description of diesel exhaust odor that reflects aldehydic and other highly odorous, somewhat fragrant components. The intensity series that consists of these pervasive odorants is represented by comparatively dilute solutions.

Composition of A

Component	Percent by weight
n-Butylbenzene	16.6
sec-Butylbenzene	16.6
p-Cymene	16.6
Heptaldehyde	38.9
Nonaldehyde	9.7
Salicylaldehyde	0.4
Cinnamic aldehyde	0.4
alpha-Methylcinnamic aldehyde	0.4
p-Tolyl aldehyde	0.4
	100.

Intensity series (diluent: mineral oil)

		Concentration, A/(A + diluent)				
Odor intensity	Code	Expon.	Fract.	Decimal	%	ppm
Slight	A-1	2^{-16}	1/65536	0.00001525	0.0015	15
Moderate	A-2	2^{-14}	1/16384	0.0000610	0.0061	61
Strong	A-3	2^{-12}	1/4096	0.000244	0.0244	244
Extreme	A-4	2^{-10}	1/1024	0.000976	0.0976	976

Preparation

Make up the stock solution A. The components are highly odorous and should be handled in the fume hood.

To make A-4, calibrate a pipet or dropper in mg/drop of stock solution A. Pipet a measured quantity of A into a flask. Add mineral oil in the ratio 1.02 g mineral oil/mg A. Shake the mixture to dissolve the components.

To make A-3, mix 1 part of A-4 with 3 parts of diluent.

To make A-2, mix 1 part of A-3 with 3 parts of diluent.

To make A-1, mix 1 part of A-2 with 3 parts of diluent.

Note: (a) The dilutions can be made on a volume basis without introducing any significant error.

MASKING (M)

Odor-modifying agents are designed to improve objectionable odors by admixture with another vapor that will change the malodor quality (masking action) and/or reduce the malodor intensity (odor-counteracting action). Agents of this type differ among manufacturers and are proprietary. When such agents are used for modification of diesel odor it will be helpful to have a general "masking" standard to use as a quality reference. Such a standard should contain compo-

nents of the type likely to resemble the composition of common diesel masking agents or to be associated in the experience of panel members with commercial products that contain mixtures of industrial essential oils.

Composition of M

Component	Percent by weight
Oil of wintergreen	10
Terpineol	20
Cedrene	20
Bornyl acetate	30
Phellandrene	20

Intensity series (diluent: mineral oil)

		Concentration, M/(M + diluent)				
Odor intensity	Code	Expon.	Fract.	Decimal	%	ppm
Slight	M-1	2^{-16}	1/65536	0.00001525	0.0015	15
Moderate	M-2	2^{-13}	1/8192	0.000122	0.0122	122
Strong	M-3	2^{-10}	1/1024	0.000976	0.0976	976
Extreme	M-4	2^{-7}	1/128	0.00781	0.781	

Preparation

Make up the stock solution M. The components are highly odorous and should be handled in the fume hood.

To make M-4, calibrate a pipet or dropper in mg/drop of stock solution M. Pipet a measured quantity of M into a flask. Add mineral oil in the ratio 0.127 g mineral oil/mg M. Shake the mixture to dissolve the components.

To make M-3, mix 1 part of M-4 with 7 parts of diluent.

To make M-2, mix 1 part of M-3 with 7 parts of diluent.

To make M-1, mix 1 part of M-2 with 7 parts of diluent.

Note: (a) The dilutions can be made on a volume basis without introducing any significant error.

DIESEL (D)

As stated earlier, the intensity of diesel exhaust odor may be measured without regard to the quality of the odor. The intensity reference standards, therefore, could consist of a dilution scale of any convenient odorant. It is considered likely, however, that panel members will become proficient in QI odor profile work more easily if the overall intensity reference standard is related to the individual odor quality standards.

Composition of D

Component	Percent by volume
B-4	59.3
O-4	37.0
A-4	3.7
	100.

Intensity series (diluent: mineral oil)

		Concentration, D/(D + diluent)				
Odor intensity	Code	Expon.	Fract.	Decimal	%	ppm
	D-1	2^{-11}	1/2048	0.000488	0.0488	488
Slight	D-2	2^{-10}	1/1024	0.000976	0.0976	976
	D-3	2^{-9}	1/512	0.00195	0.195	
	D-4	2^{-8}	1/256	0.00391	0.391	
Moderate	D-5	2^{-7}	1/128	0.00781	0.781	
	D-6	2^{-6}	1/64	0.0156	1.56	
	D-7	2^{-5}	1/32	0.0313	3.13	
Strong	D-8	2^{-4}	1/16	0.0625	6.25	
	D-9	2^{-3}	1/8	0.125	12.5	
	D-10	2^{-2}	1/4	0.25	25.	
Extreme	D-11	2^{-1}	1/2	0.5	50.	
	D-12	2^{-0}	1	1	100.	

Procedure

Make up the stock solution D.
D-12 is identical with stock solution D.
To make D-11, mix 1 part of D-12 with 1 part of diluent.
To make D-10, mix 1 part of D-11 with 1 part of diluent.
Continue in this manner until D-1 is prepared.

Notes: (a) The dilutions can be made on a volume basis without introducing any significant error.

(b) Stock solution B should be shaken before dilutions are made to bring into suspension any component of the oil of cade that might have settled down.

SOURCES AND PURITY OF CHEMICALS

All of the chemicals used must be pure enough that any impurities present do not make a detectable contribution to odor. For highly odorous chemicals, such as aldehydes, this requirement normally poses no problem.

A good odorless mineral oil is Primol 325 available from Humble Oil & Refining Company, Hutchinson River Parkway, Pelham, New York.

The benzyl benzoate should not have any of the cherry or almond odor that may be associated with the presence of some benzaldehyde impurity. Satisfactory grades are available from Mallinkrodt Chemical and from the Matheson Company.

The octyl benzene must not emit any of the sulfur odor that is sometimes associated with inferior samples. A satisfactory grade may be purchased from the Humphrey Chemical Company, in North Haven, Connecticut.

Acetylenic compounds are available from Farchan Research Laboratories, Willoughby, Ohio.

CONTAINERS

25 ml of each odorant reference solution is placed in a labeled 4-oz polyethylene squeeze bottle, fitted with a screw cap having a conical polyethylene liner.

The entire kit comprises 32 bottles (12 of the intensity series and 4 each of the 5 components of the QI profile, including the masking standards). These bottles should be arranged in a metal rack in a metal box that is provided with an internal activated-carbon panel to keep the atmosphere in the box odor-free. Porous materials of construction like wood or cardboard should be avoided because they absorb and retain odor. The activated carbon should be granular material of the type commonly used for air purification.

APPENDIX C: STATISTICAL DERIVATIONS

STATISTICAL DERIVATION OF TRIANGLE TEST SELECTIONS

We have a number of candidates who take **m** triangle tests each. On the basis of the number of correct answers, we wish to choose the "best" candidates. It is assumed (a) that the candidates act independently of each other, and (b) that the test has been demonstrated to them beforehand so that they are not learning during the test sequence. Each test is therefore independent of the other tests taken by the same candidate.

Now, if a candidate answers purely by guesswork, the chance of getting a correct answer on any single test is 1/3. The probability of getting **x** correct answers out of **m** is expressed by the probabilities of the binomial distribution:

$$P_x = \frac{m!}{x!\,(m-x)\,!}\,(1/3)^x\,(2/3)^{m-x}$$

where p_x = probability that a candidate would give x correct answers.

m = number of triangle tests given.

x = number of correct answers.

Suppose we have two candidates, one who has **x** correct answers and one who has **y**, and we want to know whether the difference between the two is significant. We therefore want the distribution of /x-y/. According to "distribution theory" in probability, **x** and **y** are identically, independently distributed binomial random variables, where

$$P_x = \frac{m!}{x!\,(m-x)!}\,(1/3)^x\,(2/3)^{m-x} \quad \text{and}$$

$$P_y = \frac{m!}{y!\,(m-y)!}\,(1/3)^y\,(2/3)^{m-y}$$

The expected value (E = expectation) of any function **f(x)** is defined as

$$E\,(f(x)) = \sum_{x=0}^{m} p_x f(x)$$

By convention, we choose $f(x) = t^x$, where **t** is a dummy variable.

Then,

$$E(t^x) = \sum_{x=0}^{m} t^x p_x \quad \text{and}$$

$$E(t^y) = \sum_{x=0}^{m} t^y p_y$$

} These are called the "generating functions" of p_x and p_y.

Now, we want $p_{(x-y)}$, the probability distribution of (x-y).

$$E(t^{x-y}) = E(t^x)\, E(t^{-y})$$

$$E(t^x) = \sum_{x=0}^{m} t^x p_x = \sum_{x=0}^{m} t^x \left(\frac{m!}{x!\,(m-x)!}\right) (1/3)^x\, (2/3)^{m-x}$$

$$= \sum_{x=0}^{m} \left(\frac{m!}{x!\,(m-x)!}\right) \left(\frac{1}{3}t\right)^x \left(\frac{2}{3}\right)^{m-x}$$

$$= \left(\frac{1}{3}t + \frac{2}{3}\right)^m \qquad \text{(by binominal expansion)}$$

and similarly,

$$E(t^{-y}) = \sum_{y=0}^{m} \left(\frac{m!}{y!\,(m-y)!}\right) \left(\frac{1}{3t}\right)^y \left(\frac{2}{3}\right)^{m-y}$$

$$= \left(\frac{1}{3t} + \frac{2}{3}\right)^m \qquad \text{(by binominal expansion)}$$

Then,

$$E(t^{x-y}) = \left(\frac{1}{3}t + \frac{2}{3}\right)^m \left(\frac{1}{3t} + \frac{2}{3}\right)^m$$

$$= \left[\left(\frac{1}{3}t + \frac{2}{3}\right)\left(\frac{1}{3t} + \frac{2}{3}\right)\right]^m$$

$$= \left[\frac{5}{9} + \frac{2}{9}\left(t + \frac{1}{t}\right)\right]^m$$

This is the **generating function** of $P_{(x-y)}$.

To find the distribution it is necessary to expand the generating function, gather like terms, and consider two competing candidates, **A** and **B**, whose respective scores are **x** and **y**. The probability that, by chance alone, the difference between the two scores is **z** (or that $x - y = z$) is the coefficient of the term in **t**. Now, if two candidates take **m** triangle tests each, the maximum difference between scores is $+ m$ or $- m$. The probability that, by chance alone, the difference between the scores is z, or $x - y = z$, is the coefficient of the term in t.

Let m = o (there are no tests)

Then,

$$E(t^{x-y}) = 1$$

$$= 1\,(t^o)$$

and $P^o = 1$ (Probability is 1 that score is tied)

Let m = 1 (one test is carried out)

Then,

$$E(t^{x-y}) = \frac{5}{9} + \frac{2}{9}\,t + \frac{2}{9t}$$

$$= \frac{2}{9}\,t^{-1} + \frac{5}{9}\,t^{0} + \frac{2}{9}\,t$$

and $P_{-1} = \frac{2}{9}$

$P_0 = \frac{5}{9}$

$P_1 = \frac{2}{9}$

or, the chances are $\frac{2}{9}$ that A will win over B, $\frac{2}{9}$ that B will win over A, and $\frac{5}{9}$ that A and B will tie.

Let m = 2

Then,

$$E(t^{x-y}) = [\frac{5}{9} + \frac{2}{9}\,(t + \frac{1}{t})]^2$$

$$= \frac{4}{81}\,t^{-2} + \frac{20}{81}\,t^{-1} + \frac{33}{81}\,t^{0} + \frac{20}{81}\,t + \frac{4}{81}\,t^{2}$$

Therefore,

m = 3

$P_{-3} = 8/729$
$P_{-2} = 60/729$
$P_{-1} = 174/729$
$P_0 = 245/729$
$P_1 = 174/729$
$P_2 = 60/729$
$P_3 = 8/729$

By computer calculations for

M = 4

$P_{-4} = 16/6561$
$P_{-3} = 160/6561$
$P_{-2} = 664/6561$
$P_{-1} = 1480/6561$
$P_0 = 1921/6561$
$P_1 = 1480/6561$
$P_2 = 664/6561$
$P_3 = 160/6561$
$P_4 = 16/6561$

for **m = 5**

$P_{-5} = 32/59,049$
$P_{-4} = 400/59,049$
$P_{-3} = 2,160/59,049$
$P_{-2} = 6,600/59,049$
$P_{-1} = 12,570/59,049$
$P_0 = 15,525/59,049$
$P_1 = 12,570/59,049$
$P_2 = 6,600/59,049$
$P_3 = 2,160/59,049$
$P_4 = 400/59,049$
$P_5 = 32/59,049$

Now, in a series of triangle tests taken by candidates A and B, the distribution of the differences in correct scores lx-yl tells us what the probability is that any given difference in scores obtained by A and B is due to chance alone. These distributions appear in Tables C1 and C2.

TABLE C1 — TRIANGLE TESTS, DISTRIBUTION OF DIFFERENCES IN SCORES

Number of triangle tests, m	Distribution of Plx-yl
0	$P_0 = 1$
1	$P_0 = 5/9$ $P_1 = 4/9$
2	$P_0 = 33/81$ $P_1 = 40/81$ $P_2 = 8/81$
3	$P_0 = 245/729$ $P_1 = 348/729$ $P_2 = 120/729$ $P_3 = 16/729$
4	$P_0 = 1921/6561$ $P_1 = 2960/6561$ $P_2 = 1328/6561$ $P_3 = 320/6561$ $P_4 = 32/6561$
5	$P_0 = 15525/59049$ $P_1 = 25140/59049$ $P_2 = 13200/59049$ $P_3 = 4320/59049$ $P_4 = 800/59049$ $P_5 = 64/59049$

TABLE C2 — TRIANGLE TESTS,
CUMULATIVE DISTRIBUTION FUNCTIONS

Number of triangle tests, m	Plx-yl		Probability that the difference in scores of at least x — y would be obtained by chance alone	
0	P_0	= 1	1.00	100%
1	P_0	= 1	1.00	100%
	P_1	= 4/9	0.44	44%
2	P_0	= 1	1.00	100%
	P_1	= 48/81	0.59	59%
	P_2	= 8/81	0.099	9.9%
3	P_0	= 1	1.00	100%
	P_1	= 484/729	0.66	66%
	P_2	= 136/729	0.19	19%
	P_3	= 16/729	0.022	2.2%
4	P_0	= 1	1.00	100%
	P_1	= 4640/6561	0.71	71%
	P_2	= 1680/6561	0.25	25%
	P_3	= 352/6561	0.051	5.1%
	P_4	= 32/6561	0.0049	0.49%
5	P_0	= 1	1.00	100%
	P_1	= 43524/59049	0.74	74%
	P_2	= 18384/59049	0.31	31%
	P_3	= 5184/59049	0.088	8.8%
	P_4	= 864/59049	0.015	1.5%
	P_5	= 64/59049	0.0011	0.11%

STATISTICAL DERIVATION OF INTENSITY RATING TEST SELECTIONS

ASSUMPTIONS

When two candidates replace a sample at or near the correct position, they are candidates who will be considered for favorable action and it is important to make a good distinction between them.

Candidates who replace the samples very far from the correct position are not likely to be chosen and we don't care much about distinctions between "very bad" and "extremely bad."

For example: correct position = 3

Candidate A Replaces sample in position 3
Candidate B Replaces sample in position 4
Candidate C Replaces sample in position 17
Candidate D Replaces sample in position 18

Then A's score is slightly better than B's and we want to know how reliable this difference is. Candidate C and D are both very poor and we don't really care whether one is worse than the other.

SCORING

Position of replaced bottle	Score (the higher number is the worse score)
Correct position	0
Correct position ± 1	1
Correct position ± 2	4
Correct position ± 3	9
Correct position ± 4 or more	16

Each candidate must be given the same program; they must not communicate with each other.

Test Number	Correct position of sample
1	3
2	8
3	12
4	16

DISTRIBUTION OF POSSIBLE SCORES:

Correct Answer	Possible Scores	Probability of each score
3	0	1/20
	1	2/20
	4	2/20
	9	1/20
	16	14/20

Explanation:

Bottle Position:	1	2	3	4	5	6	7	8	9	20
Deviation from correct location:	−2	−1	0	+1	+2	+3	+4	+5	+6	etc.
Score	4	1	0	1	4	9	16	16	16	16

Number of ways of getting a score of 0 is 1 out of 20.
Number of ways of getting a score of 1 is 2 out of 20.
Number of ways of getting a score of 4 is 2 out of 20.
Number of ways of getting a score of 9 is 1 out of 20.
Number of ways of getting a score of 16 is 14 out of 20.

Correct Answer	Possible Scores	Probability of each score
8, 12, or 16	0	1/20
	1	2/20
	4	2/20
	9	2/20
	16	13/20

Therefore, in the sum of the four tests, the lowest possible score (the best) is 0; the highest possible score (the worst) is 64.

Table C3 gives the probability distribution of scores in the odor intensity tests.

TABLE C3 — PROBABILITY DISTRIBUTION FUNCTION AND CUMULATIVE DISTRIBUTION FUNCTION OF DIFFERENCE IN SCORES OF TWO SUBJECTS ON ODOR-INTENSITY TEST

Difference in score	Probability of Difference	Cumulative Probability of Difference	Nominal Significance Level
0	0.0899	1.0000	
1	0.0289	0.9101	
2	0.0140	0.8812	
3	0.0467	0.8672	
4	0.0372	0.8205	
5	0.0374	0.7883	
6	0.0124	0.7459	
7	0.0754	0.7335	
8	0.0465	0.6581	
9	0.0274	0.6116	
10	0.0165	0.5842	
11	0.0191	0.5677	
12	0.0834	0.5486	
13	0.0123	0.4652	
14	0.0175	0.4529	
15	0.0822	0.4354	
16	0.0466	0.3532	
17	0.0119	0.3066	
18	0.0111	0.2947	0.30
19	0.0333	0.2836	
20	0.0176	0.2502	0.25
21	0.0096	0.2326	
22	0.0259	0.2230	
23	0.0214	0.1971	0.20
24	0.0222	0.1757	
25	0.0051	0.1534	0.15
26	0.0063	0.1483	
27	0.0292	0.1423	

TABLE C3 — (Continued)

Difference in score	Probability of Difference	Cumulative Probability of Difference	Nominal Significance Level
28	0.0156	0.1131	
29	0.0046	0.0975	0.10
30	0.0164	0.0929	
31	0.0182	0.0765	
32	0.0062	0.0583	
33	0.0021	0.0521	
34	0.0073	0.0500	0.05
35	0.0055	0.0427	
36	0.0033	0.0372	
37	0.0037	0.0339	
38	0.0040	0.0302	
39	0.0053	0.0262	
40	0.0024	0.0209	
41	0.0007	0.0185	
42	0.0038	0.0178	
43	0.0038	0.0140	
44	0.0012	0.0102	0.01
45	0.0015	0.0090	
46	0.0023	0.0075	
47	0.0013	0.0052	
48	0.0004	0.0039	
49	0.0005	0.0035	
50	0.0005	0.0030	
51	0.0004	0.0025	
52	0.0003	0.0021	
53	0.0002	0.0018	
54	0.0004	0.0016	
55	0.0003	0.0012	
56	0.0001	0.0009	0.001
57	0.0002	0.0008	
58	0.0002	0.0006	
59	0.0001	0.0004	
60	0.0001	0.0003	
61	0.0001	0.0002	
62	0.0001	0.0001	0.0001
63	0.0000	0.0000	
64	0.0000	0.0000	

STATISTICAL DERIVATION OF MULTICOMPONENT ODOR IDENTIFICATION TEST SELECTIONS

A candidate is given eight known odor standards, A, B, C, D . . . H. He is then asked to identify, in three successive tests, the components of a two-component mixture, a two-component mixture, and a four-component mixture.

TWO-COMPONENT MIXTURE

Let the two components be A and B. The chances of guessing both correctly are:

$$\frac{1}{8} \times \frac{1}{7} = \frac{1}{56}$$

chance of guessing A first; guessing B next

or

$$\frac{1}{8} \times \frac{1}{7} = \frac{1}{56}$$

guessing B first; guessing A next

$$\text{Total} = \frac{2}{56}$$

The chances of guessing only one correctly are:

$$\frac{1}{8} \times \frac{6}{7} = \frac{6}{56}$$

chance of guessing A first; guessing C, D, E, F, G, or H next

or

$$\frac{6}{8} \times \frac{1}{7} = \frac{6}{56}$$

chance of guessing C, D, E, F, G, or H first; guessing A next

Similary, chances of guessing only

$$B = 2 \times \frac{6}{56} = \frac{12}{56}$$

$$\text{Total} = \frac{24}{56}$$

The chances of getting none correct are

$$\frac{6}{8} \times \frac{5}{7} = \frac{30}{56}$$

chances of guessing C, D, E, F, G, or H first; chances of guessing C, D, E, F, G, or H next

$$\text{Grand total} \quad \frac{56}{56}$$

Three-component mixture and four-component mixture can be treated in the same way. The summaries are:

Number of components in mixture	Number of components identified correctly	Probability of identifying this number of components by chance
2	0	30/56
	1	24/56
	2	2/56
3	0	60/336
	1	180/336
	2	90/336
	3	6/336
4	0	24/1680
	1	384/1680
	2	864/1680
	3	384/1680
	4	24/1680

SCORING

It is assumed that it is much better to get a score of j correct than a score of j-1. Therefore, a quadratic scoring system is used. In other words, there is no great trick in getting 1 out of 4 right, even if chance is not involved, but a candidate who gets 4 out of 4 right is doing very much better, and we therefore credit him with being not (4/1) times better but $(4/1)^2$ times better. Therefore, the score is taken to be (the number of correct answers)^2.

The test program consists of giving each candidate three tests (the test components are the same for each candidate).

Test 1: two-component mixture

Test 2: three-component mixture

Test 3: four-component mixture

Candidates are selected on the following basis:

Let the score of candidate 1 equal X and the score of candidate 2 equal Y. Then the difference between scores is X — Y.

When is this difference significant? What is the basis for choice between candidates?

The lowest possible total score is $0 + 0 + 0 = 0$.

The highest possible total score is $2^2 + 3^2 + 4^2 = 29$; hence the maximum difference $(X - Y) = 29 - 0 = 29$.

Table C4 lists all of the probabilities of difference between scores from 0 to 29.

TABLE C4 — DISTRIBUTION FUNCTION AND CUMULATIVE DISTRIBUTION FUNCTION FOR ABSOLUTE DIFFERENCE IN SCORES OF ANY TWO SUBJECTS ON MULTICOMPONENT ODOR IDENTIFICATION TEST.

Difference in score	Probability of Difference	Cumulative Probability of Difference	Nominal Significance Level
0	0.1132	1.0000	
1	0.1518	0.8868	
2	0.1025	0.7350	
3	0.1397	0.6325	
4	0.1278	0.4928	
5	0.0938	0.3650	
6	0.0647	0.2712	0.25
7	0.0511	0.2062	0.20
8	0.0570	0.1554	0.15
9	0.0370	0.0984	0.10
10	0.0139	0.0614	
11	0.0148	0.0475	0.05
12	0.0141	0.0327	
13	0.0062	0.0186	
14	0.0032	0.0124	
15	0.0037	0.0092	0.01
16	0.0027	0.0055	
17	0.0011	0.0028	
18	0.0008	0.0017	
19	0.0006	0.0009	0.001
20	0.0002	0.0003	
21	0.0001	0.0001	0.0001
22	0.0000	0.0000	
23	0.0000	0.0000	
24	0.0000	0.0000	
25	0.0000	0.0000	
26	0.0000	0.0000	
27	0.0000	0.0000	
28	0.0000	0.0000	
29	0.0000	0.0000	